Allianz

Alternative Risk Transfer – at Allianz Commercial

UPDATE AND CASE STUDIES

Anneliese Stonebridge

Today's session

1 ART @ Allianz Update

2 What is ART Structured Insurance – a quick reminder

3 What are we seeing from the market?
- Client requests and Case Studies

4 ART: the future

ART @ Allianz update

Allianz

New Global Head appointed in March: Lara Martiner

Restructure has been taking place across the regions to accommodate growing UW team and geographical spread inc APAC

Continue to be challenged with large global growth targets 35%+
- largest of any LoB @ AGCS

Huge level of interest still in ART
- Broker RFP's now all include elements of ART/Out of the Box thinking
- New competitors and lots of interest from other insurers to enter the market

In short despite softening markets globally, interest in ART and structured insured is still growing and should be considered a key part of corporate insurance strategy

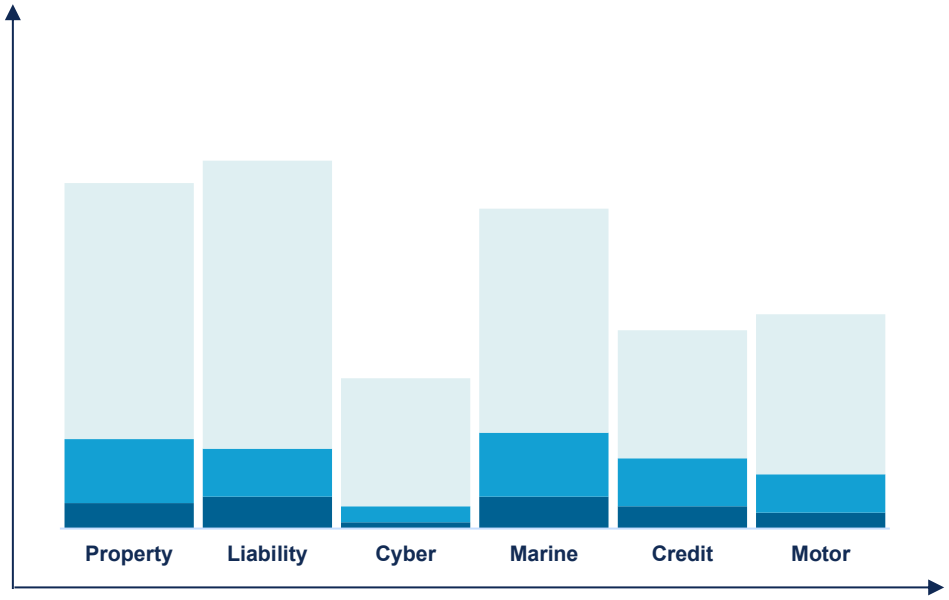
Allianz

The 4 Main Characteristics of an ART Structured Solution

1. Retention	Client retains more to save more (up to a point)
2. Sideways Risk	Leads to increased frequency risk and volatility
3. Volatility Protection	How does the client want to protect this risk/retention?
4. Diversification	Can you increase **c**ost efficiencies by including more lines (ML) and over multiple years (MY)

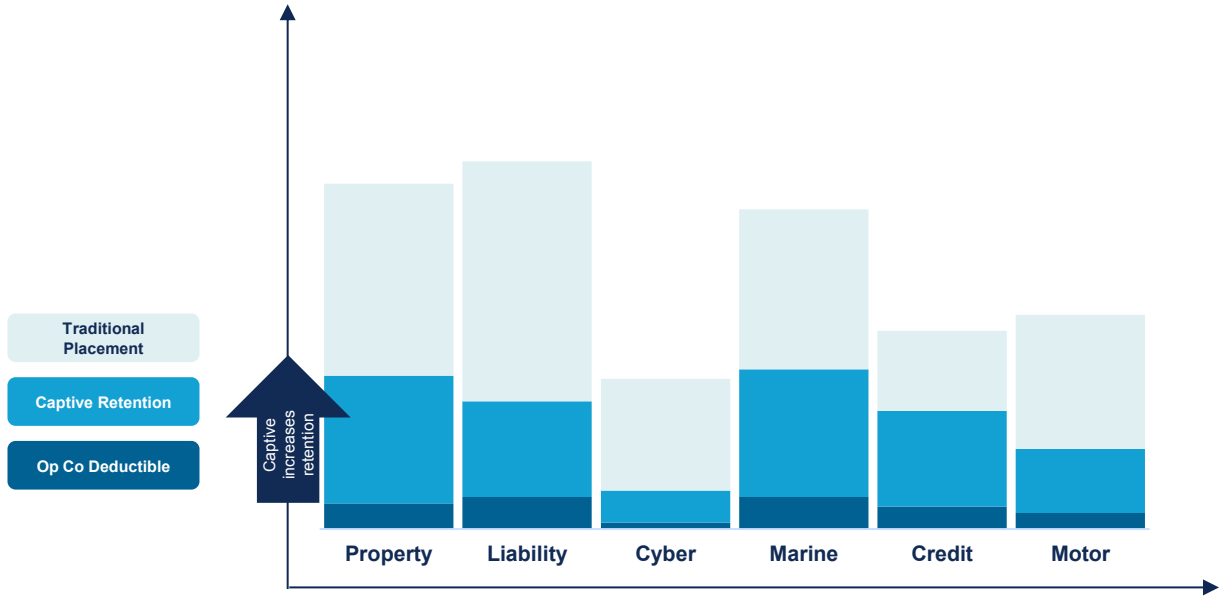
How does it work in practice?

1. Traditional placement structure



1. Captive retention is 20mn across all the lines

2. Captive vehicle increases retention



1. Captive retention is doubled to 40mn in total across all lines
2. This pushes up the attachment point at which the captive buys traditional lines, resulting in savings
3. But what if the captive wishes to keep its balance sheet retention at 20mn; what options are available to protect the 20xs20?
4. Stop loss, swing solution, multi-year multi-line (MYML)…

What are we seeing?

- US Casualty and US Auto Liability still a major focus
- Overall Downside Protection – consolidated sideways exposure management
- Much larger retentions – volatility management
- ART supporting other areas of Allianz to write more
  - Fronting for other LoBs
- Uninsurable risk – consolidation of business/operational risk into the captive
- Medical and Employee Benefits in the Captive
- Cyber – to buy or not to buy? That is the question…
- Parametric – How can we be different?

US Auto Liability

Structured Auto Liability deal for a transportation and logistics company

Background

A large notable trucking carrier, offers freight solutions across North America. Auto Liability cover is crucial for the clients operations and the traditional market capacity continues to deteriorate year on year, with prices rising.

Challenges

The US Auto Liability market has observed significant rate increases over the past few years owing to significant claims inflation and ever-increasing jury nuclear verdicts. In most cases, cover is unaffordable or simply not available within the traditional market below a certain level, leaving clients carrying huge uninsured per occurrence retentions and exposing the balance sheet to increasing levels of volatility in the event of multiple large claims.

Solution

ART wrote a multi-year 'swing' deal across multiple layers of insurance, to provide the insured with a cost-effective alternative that ensures a strong alignment of interest between the insured and insurer as the insured has a lot of ‘skin in the game’ prior to risk transfer.

Results

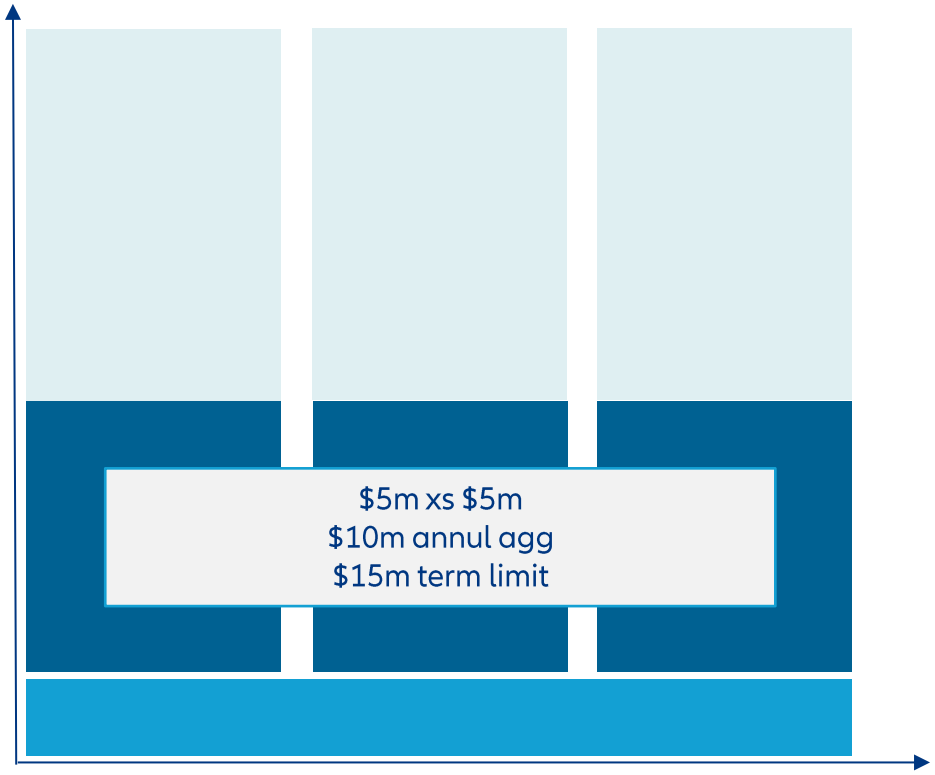
- **Offered flexible line sizes that fit within the client’s budget while still delivering the required protection**
- **Continuity of coverage for client enabling then to operate**
- Delivered a multi-year solution that offered pricing stability over time, aligning with client's long-term risk management strategy

Allianz

US Auto Liability Deal Overview

Sector	Transport
Term	Three Year
Coverage	Structured Auto Liability
Limits of Liability	$5m Per Occurrence $10m Annual Agg $15m Term Limit Excess of $3 xs $2m Per Occurrence
Deposit Premium	$4m over the term
Return Premium	25% of premium less claims if < $3m with a No Claims Bonus of $750K
Additional Premium	AP provision in the event losses exceed $3m up to a maximum of $6m
Reinstatement	Additional limit available priced at inception

Structure



*Numbers adjusted for example purposes

Energy Distribution

Multi Year Multi Line Stop Loss

Background

The client is an energy distribution company that primarily focuses on transporting and distributing energy via their own infrastrcture. Over the years, they have increased their retentions across their insurance program, allowing them to utilise their captive more efficiently and give opportunity to the captive to provide tailored cover to the client which might not be available in the traditional market.

Challenges

The initial challenge was the size of the required increase in captive retention across all their major insurance programs. Given this material increase, the client wanted to have a stop loss solution in place should the captive be eroded by multiple catastrophic events or an accumulation of smaller events eroding the net retention.

Solution

To address the increase in exposure for the captive, ART led a multi line stop loss solution, to cover all the captive exposure. As a result, the client has the peace of mind that the captive has adequate risk transfer protection plus capital relief (requirement of the local regulator) in the event of large and multiple aggregated events.

Results

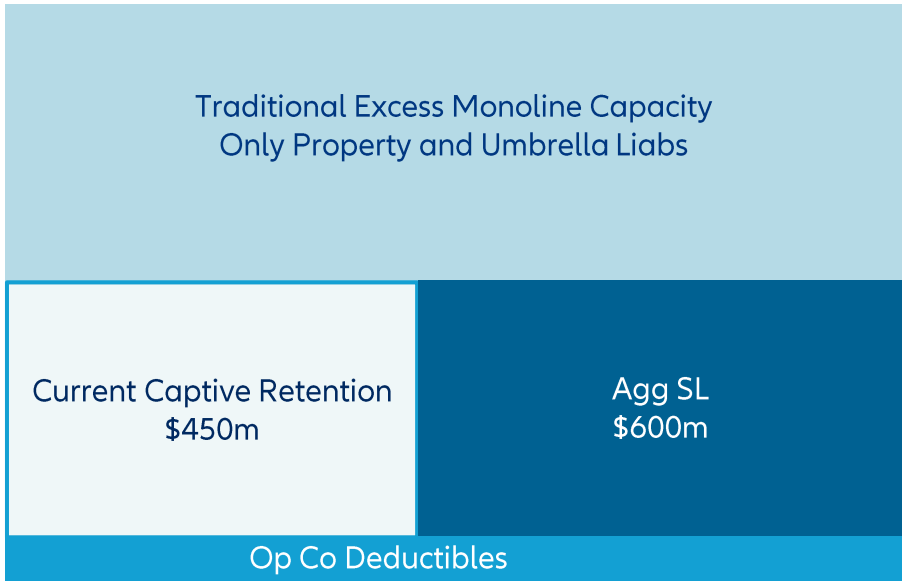
- **Captive has greater protection over catastrophic events**
- **ART fronting for traditional lines has allowed for increased capacity**
- **The client has guaranteed capacity over the long term as it is a multiyear transaction**
- **The increase in captive retention has given the client savings by pushing up the traditional layers**

Allianz

Deal Overview

Sector	Energy
Term	Three Years
Limits of Liability	$600m over the term
Annual Aggregate Deductible	$450m resets each year
Contributing Lines	Property Umbrella Liability Multiple JV Liability's 'Others' inc US Auto, Crime etc
Special Features	ART front for AGCS traditional lines in order to put out greater capacity as Allianz

Structure



*Numbers adjusted for example purposes

Hospitality

Annual Multi Line Stop Loss

Background

The insured is an international hospitality client that operates predominantly on a franchise model with minimal assets. They have a very sophisticated insurance programme with strong captive use, however, they want to take more operational risk out of the business. This risk is not insured in the traditional market but the captive also does not want to take on uncapped exposure.

Challenges

The initial challenge was understanding the exact uninsurable exposure the client wanted to cover within the captive, whether this was a buy back of an exclusion or a totally separate line of cover. Taking on this risk increased the captive exposure beyond which the client was comfortable with, so they wanted to put a stop loss solution in place to cap their overall risk in any one year, both traditional and uninsurable.

Solution

To address the increase in exposure for the captive, ART provide a multi line stop loss solution, to cover all the captive exposure including the uninsurable risk. As a result, the client has the ability to diversify the captive through mixing traditional and uninsurable risk but with a cap on the overall annual cost to the business.

Results

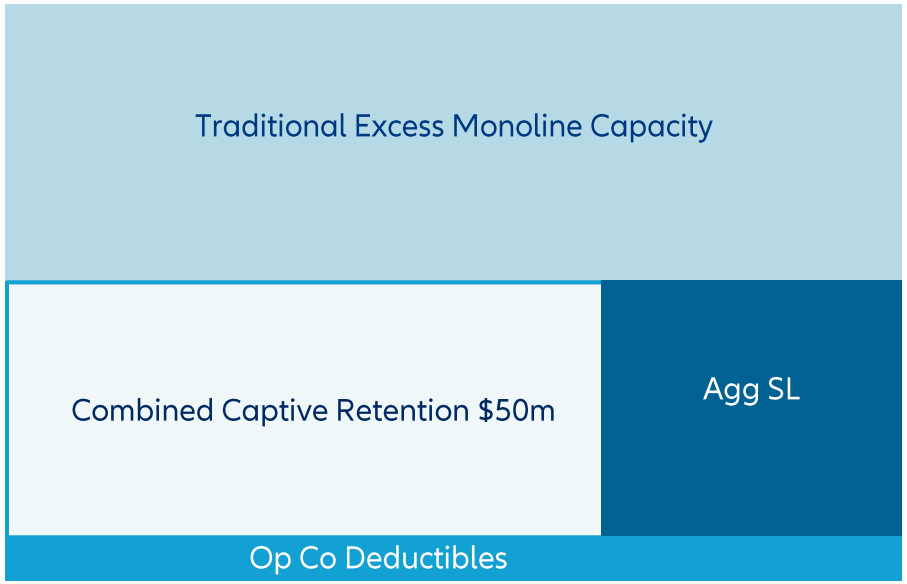
- **Captive has capped its overall risk per annum**
- **Captive can take on other uninsurable risk from the business and still protect it**
- **Bespoke wordings to support client needs for uninsurable risks**
- **Consolidates traditional, short tail EB and uninsurable risks**

Allianz

Deal Overview

Sector	Hospitality
Term	One Year
Limits of Liability	$various per line
Annual Aggregate Deductible	$50m – combined across all 25 lines
Contributing Lines	Traditional inc. Property, General Liability, Cyber Employee Benefits (Short Tail) inc. Medical, Uninsurable risk inc buy backs of Communicable Disease exclusions, operational risk etc

Structure



*Numbers adjusted for example purposes

Cyber

In the wake of recent attack's we have seen a lot of talk about the 'cost' of Cyber

ART integrate Cyber into a lot of our multi-line programmes with large corporates taking on average the first $15M into their captive

However, for those on the fence, or don't buy Cyber, could ART be the answer?

- Fully unbundled captive fronted policies – to allow admitted claims payments when necessary
- Structured mono-line or multi line captive primary layer, building cheaper excess capacity on top
  - Reduces the overall insurance cost but still provides high limit protection for catastrophic scenarios
  - Longer term policies build up funds for future losses

Parametric

Is not a new concept

Is known for being expensive with the potential for large basis risk if not done correctly

How could things be done differently?

- Structured Parametric
- Alignment of natural disasters and NDBI losses in particular, retail shut down

ART: The future is bright

The future of ART is bright:
- Growth up
- Innovation still key
- Recruitment drive across the market

Building on our global success and demands of our clients we have committed to a reinsurance treaty to expand our line size and capacity, cementing our place further as the market leader

Being more connected to other parts of the business; AGCS Traditional lines, Allianz Trade, Allianz Partners has created more opportunities for us to be innovative and support our clients

As certain parts of the market soften we will always continue to look for other ways to diversify our portfolio

Closed-door Feedback

Allianz

Performance & Uncertainty

Nathalie Nahai

RISKS, REWARDS & REIMAGINING WORK

IN THE AGE OF AI

Q

What is your greatest hope & fear for AI?

You may have seen the headlines…

SUCCESS· ARTIFICIAL INTELLIGENCE

AI expert says it's 'not a question' that AI will take over all jobs—but people will have 80 hours a week of free time

BY **EMMA BURLEIGH**
REPORTER, SUCCESS

September 9, 2025 at 10:55 AM EDT

Dr. Roman Yampolskiy, a professor of computer science and leading voice in AI safety, says the technology will take over 99% of jobs in the next five years.
OLGA PANKOVA / GETTY IMAGES

- **While many CEOs are pausing hiring while they wait and see what jobs AI could replace,** one AI expert warns that nearly all roles are at risk. Professor of computer science Dr. Roman Yampolskiy predicts that 99% of work will be placed by AI and humanoid robots

mikkelwilliam—Getty Images

MAY 30, 2025 4:53 PM CET

What Happens When AI Replaces Workers?

IDEAS AI

by Luke Drago and Rudolf Laine

On Wednesday, Anthropic CEO Dario Amodei declared AI could eliminate half of all entry level white collar jobs within five years. Last week, a senior LinkedIn executive reported that AI is already starting to take jobs from new grads. In April, Fiverr's CEO made it clear: "AI is coming for your job. Heck, it's coming for my job too." Even the new Pope is warning about AI's dramatic potential to reshape our economy.

Why do they think this?

EXCLUSIVE BUSINESS

Oracle, OpenAI Sign $300 Billion Cloud Deal

The majority of new revenue revealed by Oracle will come from OpenAI deal, sources say

By Berber Jin Follow

Updated Sept. 10, 2025 4:08 pm ET

Share Resize 109 Listen (2 min)

Most Popular News

Toxic Fumes Are Leaking Into Airplanes, Sickening Crews and Passengers

A Strange Gas-Pumping Defect Is Making $100,000 Corvettes Go Up in Flames

On the Fence About a Spending Decision? Try the 0.01% Rule

Investigators Learn More About Suspect in Kirk Shooting

'It Will Scar This Generation.' Charlie Kirk's Death Ignites a

But what *is* AI?

AI is technology that enables computers and machines to simulate human learning, comprehension, problem solving, decision making, creativity and autonomy.

IBM

https://www.ibm.com/topics/artificial-intelligence

It’s been a winding road

1940

Enigma machine was decoded using AI during World War II.

1950

Alan **Turing** releases test to assess machine intelligence.

1955

John McCarthy coins term "**AI**".

1961

Introduction of **Unimate**, the first industrial robot.

1964

Joseph Weizenbaum invents first chatbot, **ELIZA**.

1969

Shakey The Robot, the first general-purpose mobile robot, is introduced.

1995

Alice The Chatbot is introduced by Richard Wallace.

1997

DeepBlue, IBM's AI-powered chess system, defeats chess champion Garry Kasparov.

1998

Cynthia Breazeal develops **Kismet** (MIT AI Lab), a robot designed for social interactions with humans.

2000

Roomba, AI-powered vacuum cleaner, goes to market.

2008

Voice recognition on the iPhone and the birth of **Siri**.

2011

IBM Watson, the question-answering AI, is introduced.

2014

Alexa, Amazon's virtual assistant, becomes a key tool on Amazon devices.

2016

AlphaGo (Google DeepMind) beats grandmaster Lee Sedol 4-1, the first time a computer beats a top human player at Go.

2017

Sophia, a humanoid robot created by Hanson Robotics, is granted Saudi Arabian citizenship.

Taryn Southern uses **Amper** AI to compose music for album.

2020

GPT-3, a significant predecessor to ChatGPT, demonstrates advanced language understanding and text generation capabilities.

DALL-E, a neural network, is used to create high-quality images.

2021

UNESCO adopts the **Recommendation on the Ethics of AI**, the first global standard, applicable to all 194 member states of UNESCO.

2022

ChatGPT is released publicly as a free research project using the GPT-3.5 model.

2024

EU AI act is unanimously approved by the EU Council, to regulate providers of AI systems and entities using AI in a professional context.

2025

DeepSeek's R1 model emerges as a significant competitor to OpenAI's o1 model.

USA + UK refuse to sign **AI Action Summit declaration** calling for policies ensuring AI is open, inclusive, transparent, ethical, safe, secure, and trustworthy.

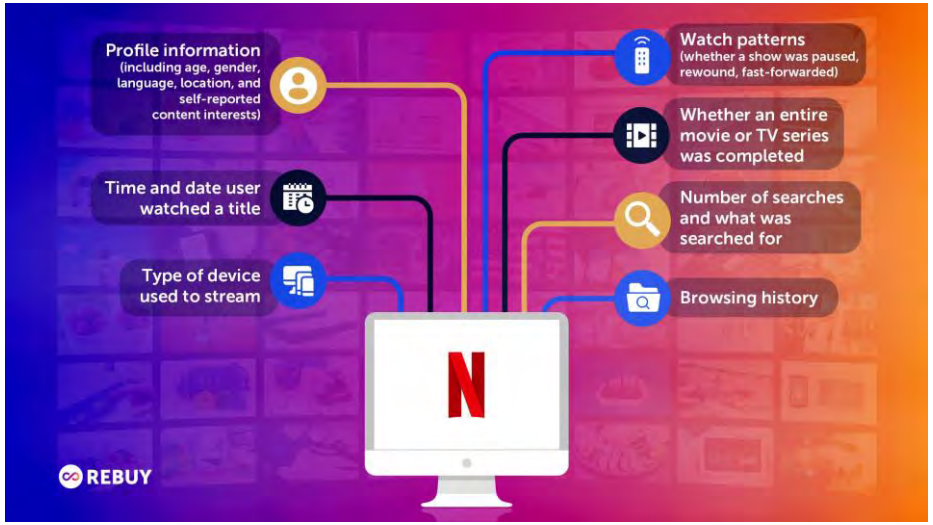
1

Types of AI

Narrow AI

Only type of AI that exists today.
Trained to perform single / narrow task, targets single subset of cognitive abilities, advances in that spectrum.

Reactive machine AI



Limited memory AI

Gemini

Copilot

SUNO

perplexity

Claude

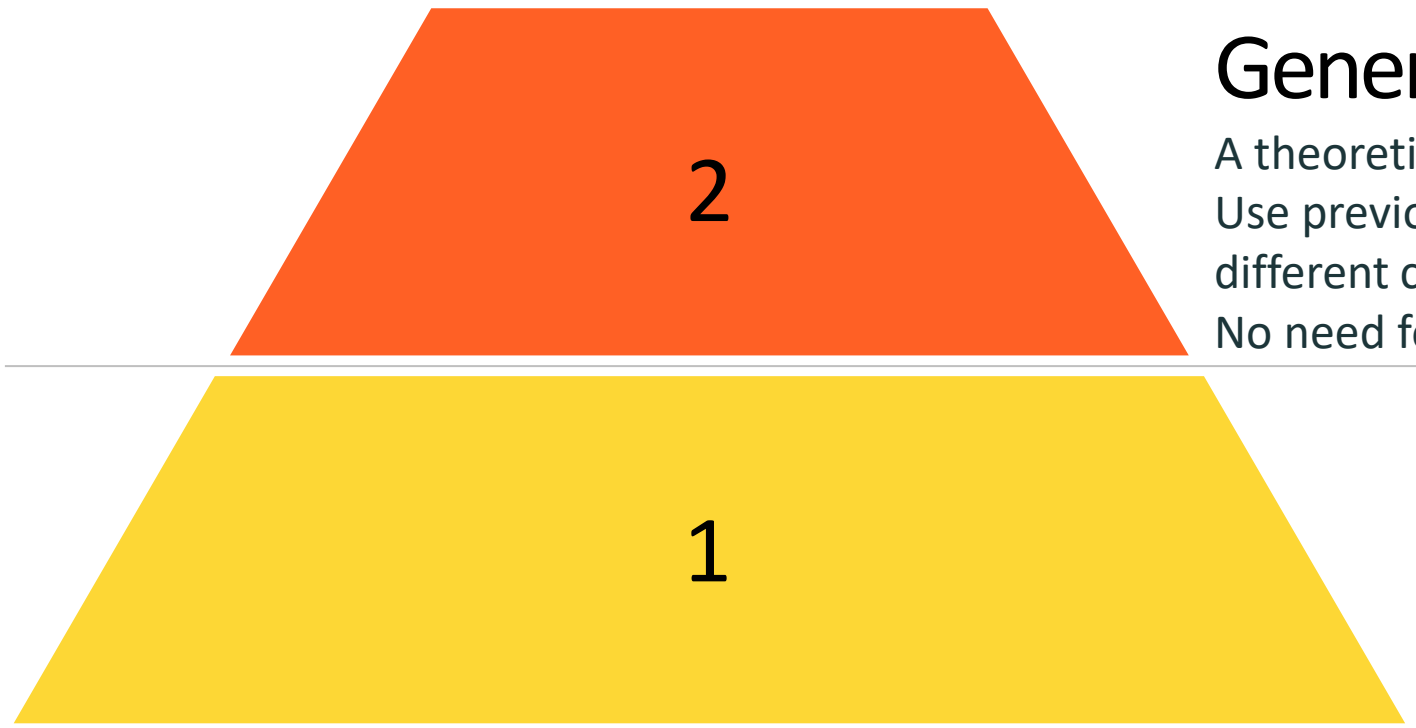
Midjourney

deepseek

ChatGPT

alexa

Siri



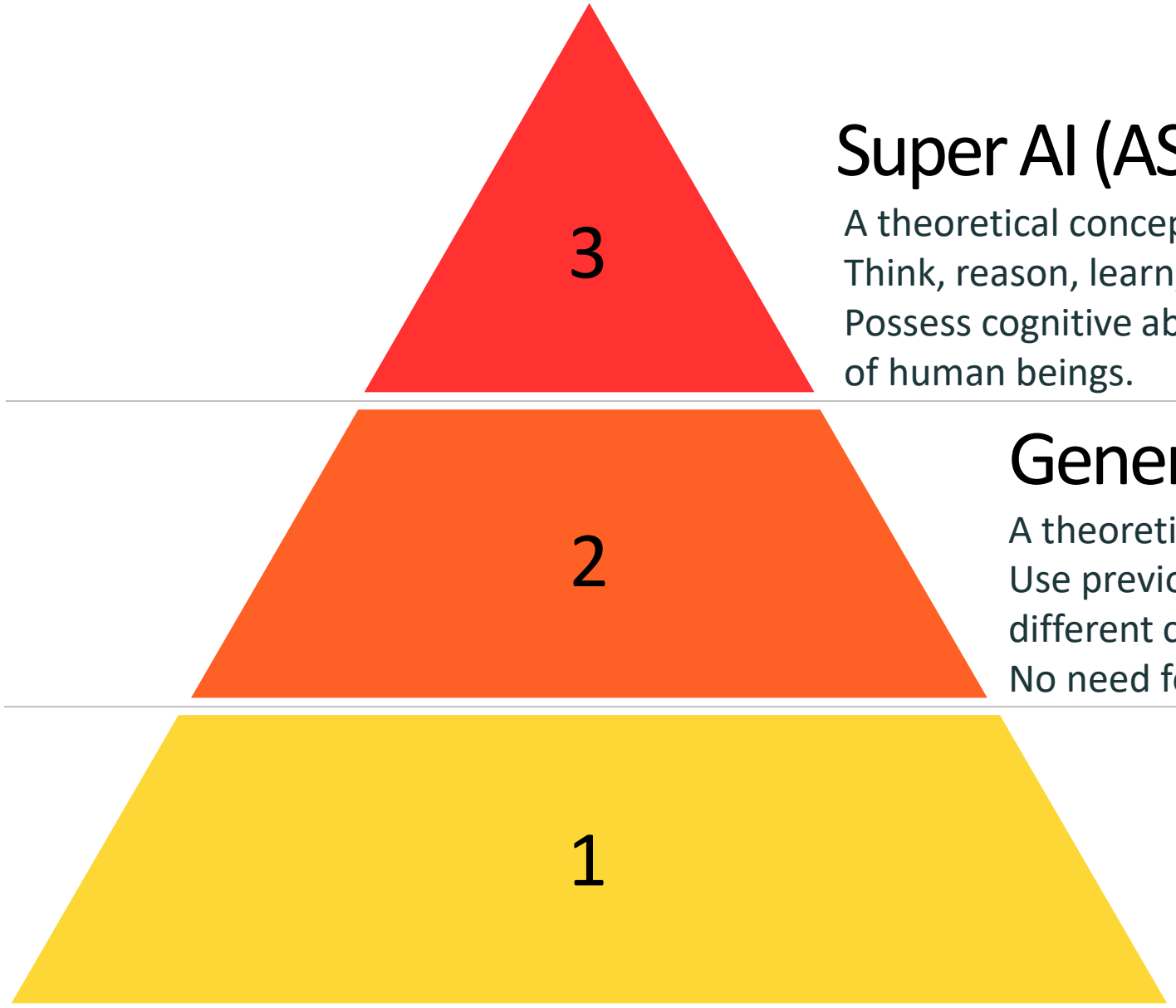
General AI (AGI)

A theoretical concept
Use previous skills + learnings to accomplish new tasks in a different context.
No need for humans to train underlying models.

Narrow AI

Only type of AI that exists today.
Trained to perform single / narrow task, targets single subset of cognitive abilities, advances in that spectrum.


ROBOTIQ



Super AI (ASI)

A theoretical concept
Think, reason, learn, make judgements.
Possess cognitive abilities that surpass those of human beings.

General AI (AGI)

A theoretical concept
Use previous skills + learnings to accomplish new tasks in a different context.
No need for humans to train underlying models.

Narrow AI

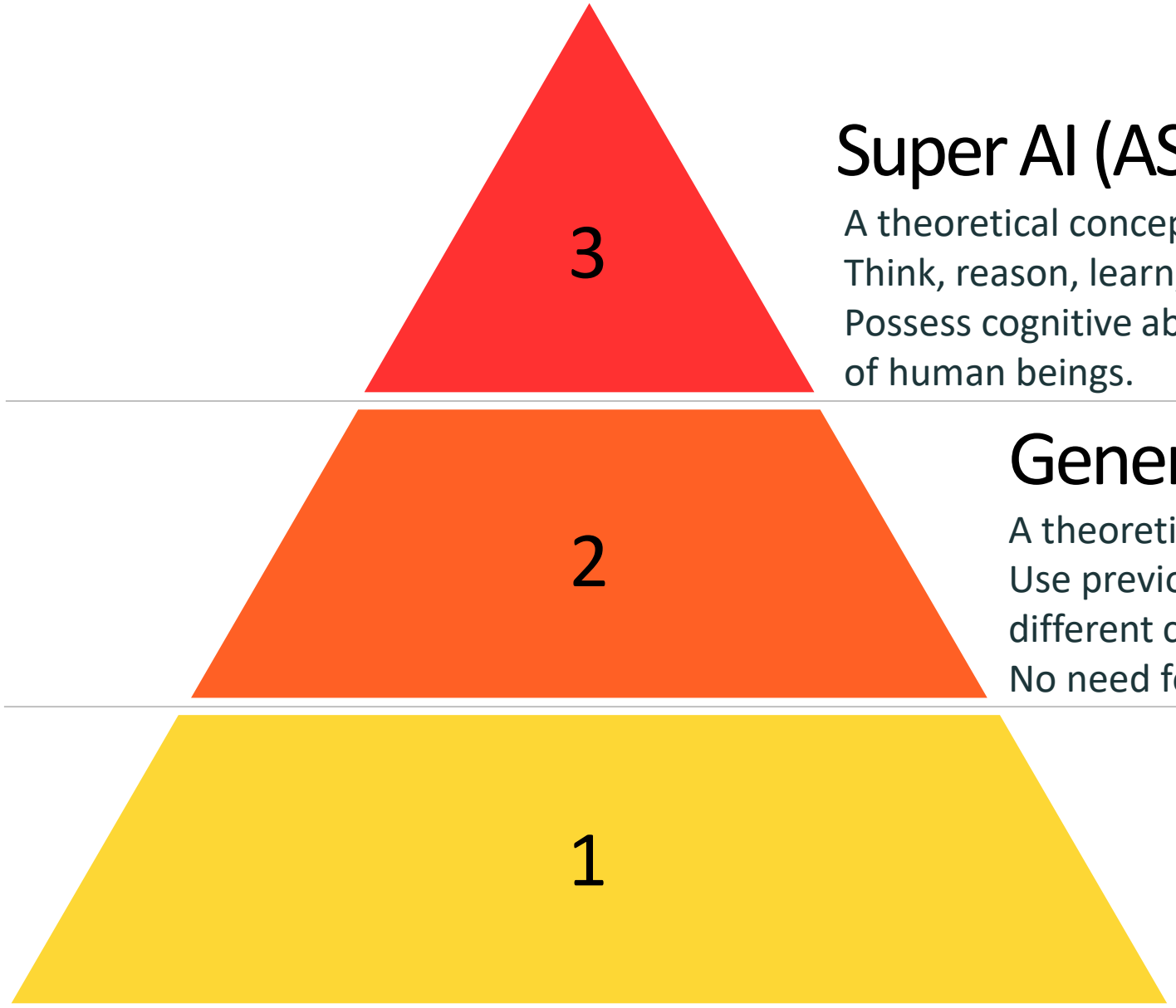
Only type of AI that exists today.
Trained to perform single / narrow task, targets single subset of cognitive abilities, advances in that spectrum.

DATA (STAR TREK – 1987-1994)

HER (SPIKE JONZE - 2013)

Q

Where do you think we are?



Super AI (ASI)

A theoretical concept
Think, reason, learn, make judgements.
Possess cognitive abilities that surpass those of human beings.

General AI (AGI)

A theoretical concept
Use previous skills + learnings to accomplish new tasks in a different context.
No need for humans to train underlying models.

Narrow AI

Only type of AI that exists today.
Trained to perform single / narrow task, targets single subset of cognitive abilities, advances in that spectrum.

ARTIFICIAL INTELLIGENCE

GPT-5 is here. Now what?

The much-hyped release makes several enhancements to the ChatGPT user experience. But it's still far short of AGI.

By Grace Huckins

August 7, 2025

Outsized Hype About ChatGPT GPT-5 Being Totally Honest AI Is Basically A Big Fib

By Lance Eliot, Contributor. Dr. Lance B. Eliot is a world-renowned AI scienti...

Published Aug 24, 2025, 03:15am EDT

AI is getting better at telling the truth but still persists as a deceptive liar.

GETTY

2025 Edelman Trust Barometer

Trust and the Crisis of Grievance

With Insights for the Technology Sector

Edelman Trust Institute

25 YEARS Edelman Trust Barometer

Enthusiasm for Use of Artificial Intelligence Declines

Percent who say

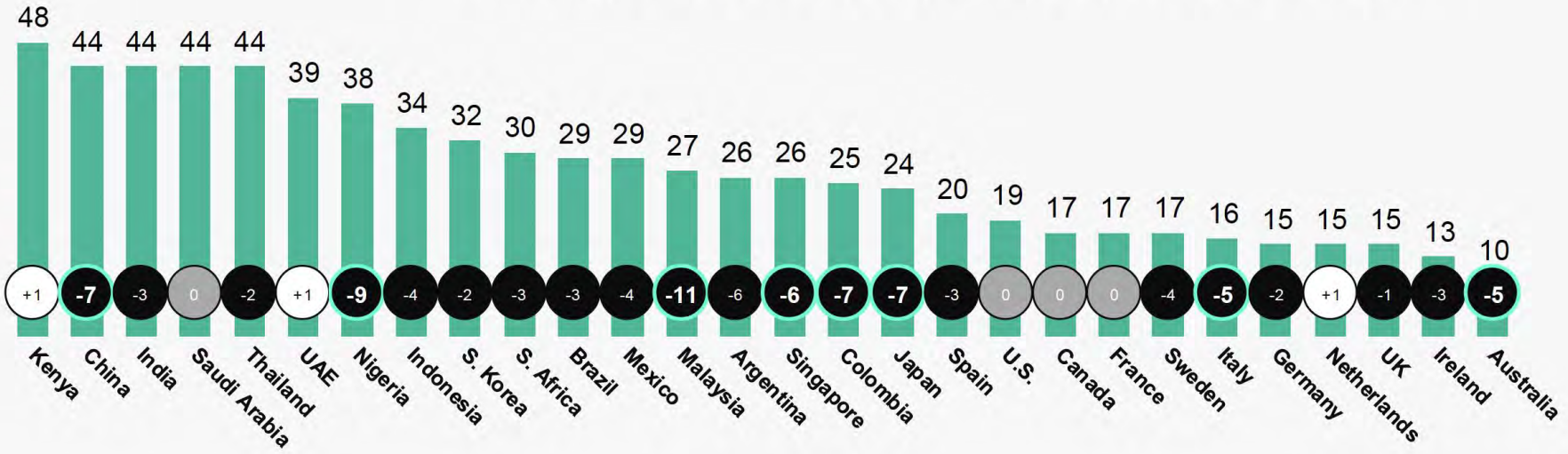
GLOBAL 28

I embrace the growing use of AI

Significant declines in
China, Nigeria, Malaysia, Singapore, Colombia, Japan, Italy, and Australia

27%

-3 pts
Change, 2024 to 2025



2025 Edelman Trust Barometer. CHG_TEC_COM. How would you characterize your feelings about each of the following? 5-point scale; bottom 2 box, resistant or hesitant = reject; top 2 box, enthusiastic or passionate = embrace.
General population, 28-mkt avg. Year-over-year changes were tested for significance using a t-test set at the 99%+ confidence level.

Nearly 1 in 2 Skeptical of Business Use of Artificial Intelligence

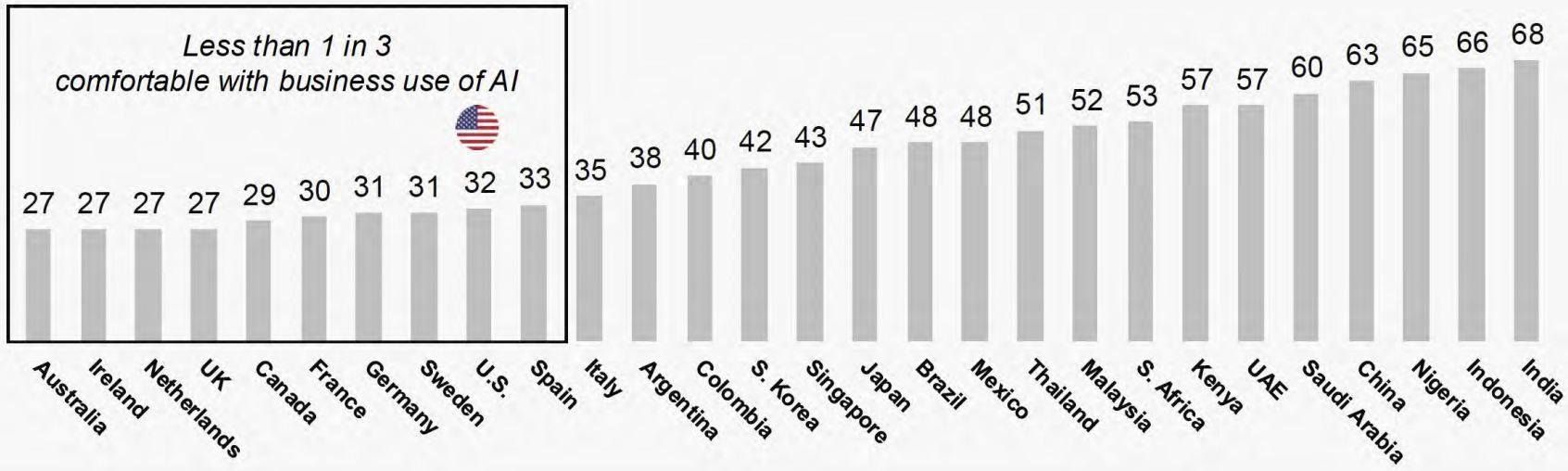
Percent who say

GLOBAL 28

I am comfortable with business using AI

ONLY

44%



2025 Edelman Trust Barometer. AI_INT_KPI. For each of the following institutions, please indicate whether their general use of artificial intelligence such as machine learning, natural language processing, and generative AI is something you are comfortable with or something you are concerned about. 5-pt scale; top 2 box, comfortable. Question asked of half the sample. General population, 28mkt avg.

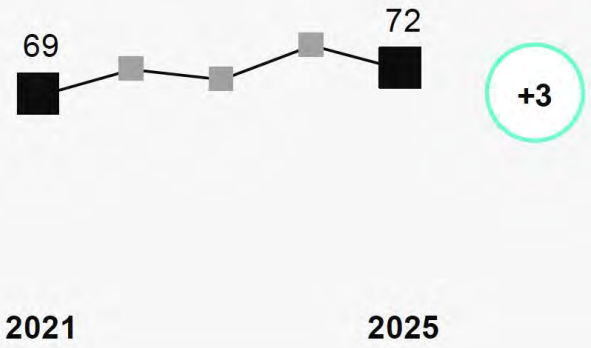
Tech-Related Societal Fears Have Increased Since 2021

Percent who say

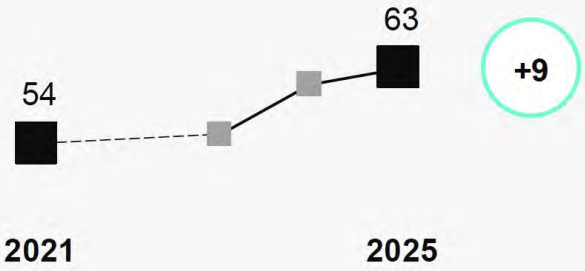
GLOBAL 26 — (-) (0) (+) Change, 2024 to 2025 — Significant change

I worry about…

Cyber Attacks

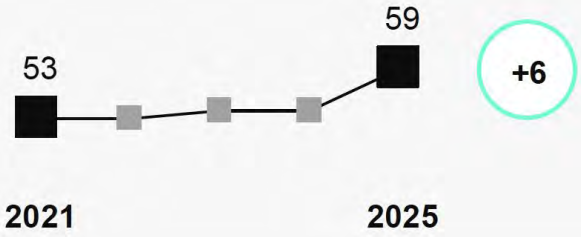


Information War



Job Loss to Automation

(Among employees)



2025 Edelman Trust Barometer. POP_EMO. Some people say they worry about many things while others say they have few concerns. We are interested in what you worry about. Specifically, how much do you worry about each of the following? 9-point scale; top 4 box, worry. "Job loss due to automation" only asked of those who are an employee of an organization (Q43/1). Some attributes only asked to half the sample. General population, 26-mkt avg. Year-over-year changes were tested for significance using a t-test set at the 99%+ confidence level.

Anker's coin-size AI recorder can transcribe and summarize your meetings - in one click

The Soundcore Work summarizes meetings, transcribes over 100 languages, and still manages to be smaller than rival devices.

Written by **Artie Beaty,** Contributing Writer

Sept. 4, 2025 at 10:55 a.m. PT

/ must read

I used a $20 AI tool to finish 24 days of coding in 12 hours

Read now →

/ related

7 most exciting tech accessories from IFA 2025 (and that you can actually buy)

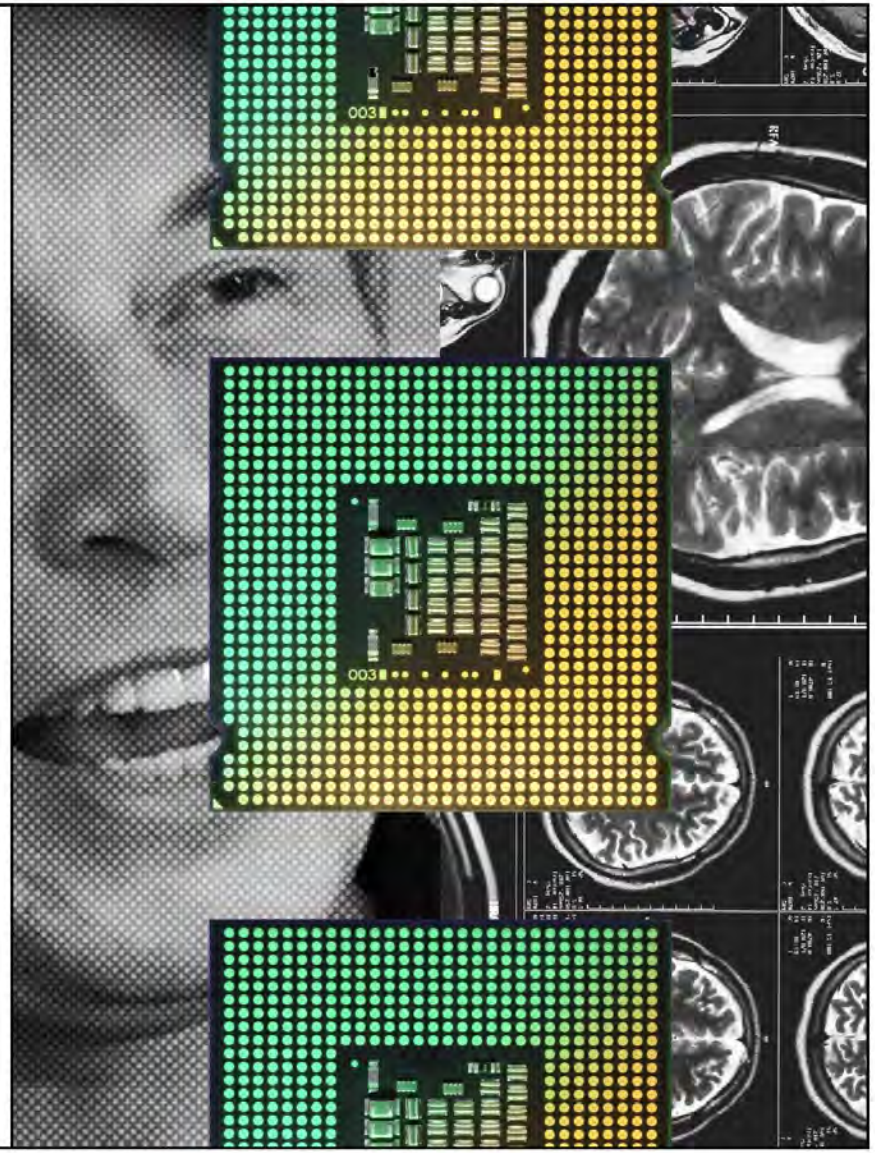
RS REPORTS

NEURALINK, ELON MUSK, AND THE RACE TO PUT CHIPS INTO OUR BRAINS

Though brain chips are all over TV and the news now with Neuralink, scientists like those at Caltech have been working on the technology for decades. And some question Musk's approach

By DARREN LOUCAIDES
Illustration By MATTHEW COOLEY

SEPTEMBER 13, 2025



There's a 3D hand on a black screen. J. Galen Buckwalter concentrates

PHOTOGRAPHS IN ILLUSTRATION BY JIM

PRESS RELEASES | SPECIAL PROCEDURES

UN expert calls for regulation of neurotechnologies to protect right to privacy

12 March 2025

Share

In the world of finance…

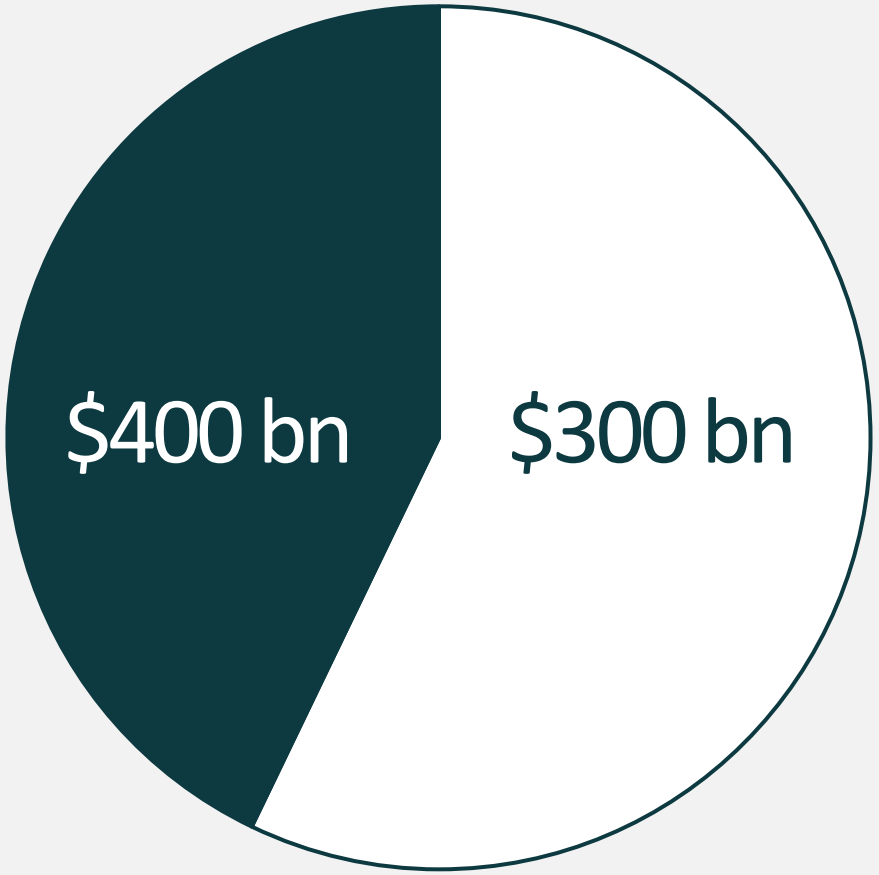
AI technologies could add up to $1.1 trillion in annual value for the global insurance industry.

McKINSEY
The Executive's AI Playbook (2023)

https://www.mckinsey.com/industries/financial-services/our-insights/insurer-of-the-future-are-asian-insurers-keeping-up-with-ai-advances

ANNUAL VALUE

Pricing, underwriting, & promotion tech upgrades



AI-powered customer service & personalised offerings

AI & FINANCE

75%

UK financial firms are already using AI

AI & FINANCE

75%

UK financial firms are already using AI

10%

More financial firms planning to use AI over the next 3 years

AI & FINANCE

75%

UK financial firms are already using AI

10%

More financial firms planning to use AI over the next 3 years

82%

Commercial insurance carriers planning agentic AI adoption within 3 years

What *is* agentic AI?

2

Agentic AI

AGENCY

Agentia (abstract noun of…)

AGENCY

Agentia	(abstract noun of…)
Agentum	“effective, powerful” (present participle of…)

AGENCY

Agentia	(abstract noun of…)
Agentum	“effective, powerful” (present participle of…)
Agere	“to set in motion, drive forward” “to do, perform; incite to action”

AGENT

One who acts

In the world of psychology…

AGENCY

Human	AI
• Control	

AGENCY

Human	AI
• Control • Autonomy	

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action	

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices	

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices • Regulate our behaviour	

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices • Regulate our behaviour • Influence life circumstances	

HEIGHTENED AGENCY

Human

Wellbeing

Creativity

Motivation

Goal achievement

In the world of technology…

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices • Regulate our behaviour • Influence life circumstances	• Agent = system or program

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices • Regulate our behaviour • Influence life circumstances	• Agent = system or program • Autonomously performs tasks

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices • Regulate our behaviour • Influence life circumstances	• Agent = system or program • Autonomously performs tasks • Designing own workflow + utilising available tools

AGENCY

Human	AI
• Control • Autonomy • Capacity for self-directed action • Ability to make choices • Regulate our behaviour • Influence life circumstances	• Agent = system or program • Autonomously performs tasks • Designing own workflow + utilising available tools • On behalf of a user / another system

Autonomous systems designed to analyse data, identify patterns, and execute tasks with minimal or no human intervention.

DELOITTE

AI-driven Transformation in the Commercial Insurance Industry

https://www.deloitte.com/us/en/Industries/financial-services/articles/commercial-insurance-industry-ai-driven-transformation.html

HEIGHTENED AGENCY

Human

Wellbeing

Creativity

Motivation

Goal achievement

AI

Make decisions

Solve problems

Execute actions

Interact with external environments

If we’re not careful…

AI

Agents

AI

Agents

Skills

Discernment

Agency

Critical thinking

AI

Agents

Skills

Discernment

Agency

Critical thinking

Us

OMG

Publication

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task

‹ Research

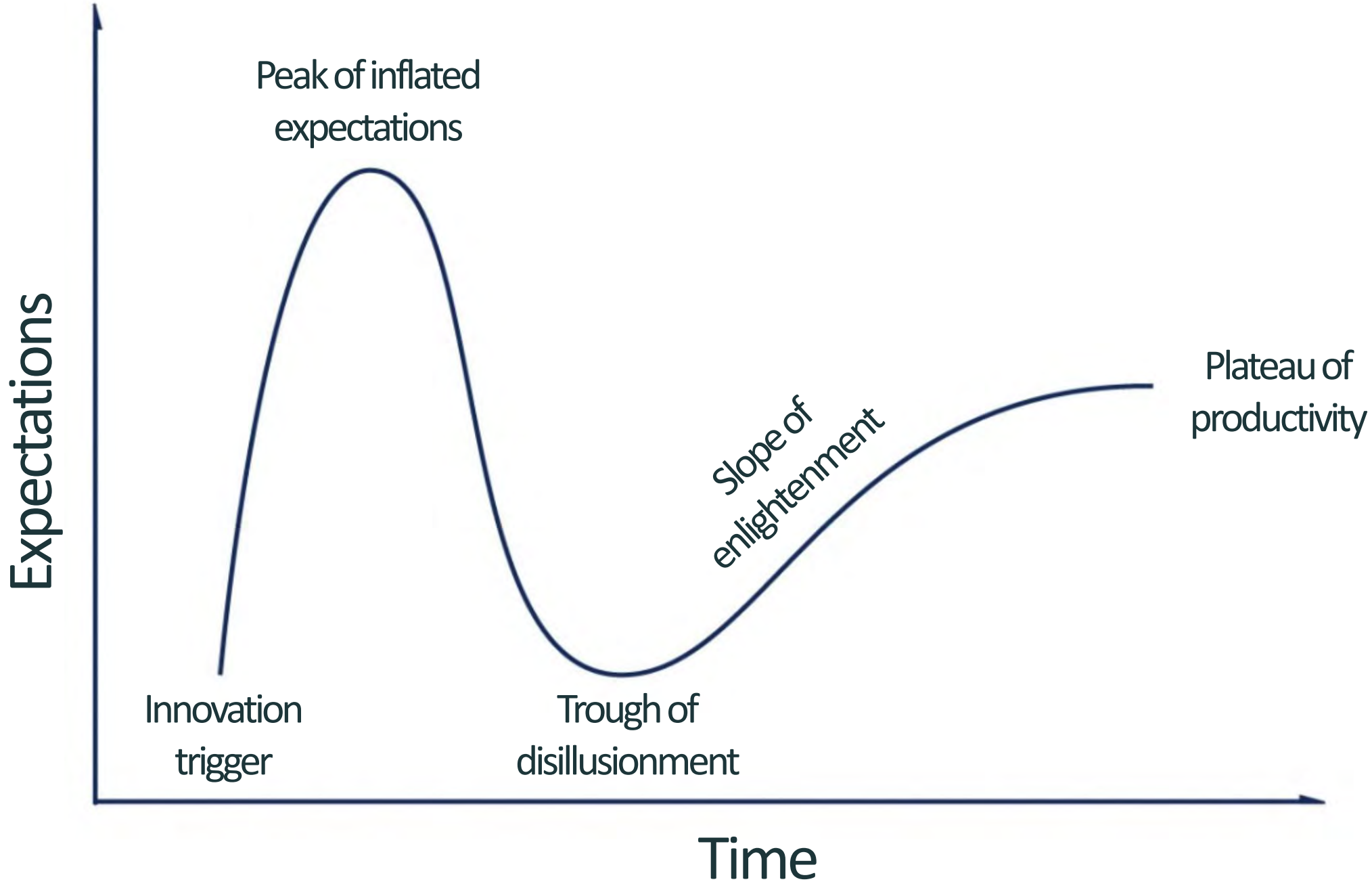
June 10, 2025

People

Nataliya Kos'myna
Research Scientist

Nataliya Kosmyna, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein, and Pattie Maes. "Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task." arXiv preprint arXiv:2506.08872 (2025).

Leveraging AI successfully requires skill & discernment


Peak of inflated expectations
Expectations
Slope of enlightenment
Plateau of productivity
Innovation trigger
Trough of disillusionment
Time

Peak bubble

It's hard to see how this won't end badly

GARY MARCUS
SEP 11, 2025

Tulip mania

Tulip mania was a period during the Dutch Golden Age when contract prices for some bulbs of the recently introduced and fashionable tulip reached extraordinarily high levels. The major acceleration started in 1634 and then dramatically collapsed in February 1637. It is generally considered to have been the first recorded speculative bubble or asset bubble in history.
Wikipedia

I don't know when the GenAI bubble will end. But this has got to be peak bubble:

> The deal is one of the largest cloud contracts ever signed, reflecting how spending on AI data centers is hitting new highs despite mounting concerns over a potential bubble.
>
> The Oracle contract will require 4.5 gigawatts of power capacity, roughly comparable to the electricity produced by more than two Hoover Dams or the amount consumed by about four million homes.
>
> Oracle shares surged by as much as 43% on Wednesday after the cloud company revealed it added $317 billion in future contract revenue during its latest quarter that ended in Aug. 31. Chief Executive Safra Catz told analysts that it had signed contracts with three different customers during the quarter.
>
> The share price surge increased Oracle Chairman Larry Ellison's wealth by more than $100 billion, pushing him into the range of Elon Musk as the world's richest person, with a net worth of almost $400 billion.
>
> The OpenAI and Oracle contract, which starts in 2027, is a risky gamble for both companies. OpenAI is a money-losing startup that disclosed in June it was generating roughly $10 billion in annual revenue—less than one-fifth of the $60 billion it will have to pay on average every year. Oracle is concentrating a large chunk of its future revenue on one customer—and will likely have to take on debt to buy the AI chips needed to power the data centers.

In short

- OpenAI doesn't have $300 billion dollars
- They don't have anywhere near $300 billion dollars
- By their own (presumably optimistic) projection, they won't turn a profit until 2030.
- And all this from a company thought (or claimed) that GPT-5 was going to be tantamount to AGI (spoiler alert: it wasn't)
- For good measure Oracle doesn't have the chips they would need to fulfill the contracts, or even the cash to buy them.

I won't say that it is *all* make-believe, but, well, you do the math. (Did people

EDITORS' PICK | INNOVATION > AI

Will AI Really Take Your Job? Experts Weigh In On Dario Amodei's Prediction

By Kolawole Samuel Adebayo, Contributor. I write about the economics of AI.

Published Jun 04, 2025, 01:50pm EDT, Updated Jun 10, 2025, 02:29pm EDT

Experts break down what AI job loss really means — and why reinvention, not just replacement, is the future of work.
GETTY

Counting The Cost

Klarna made headlines in 2024 when it replaced 700 customer support agents with an AI chatbot. But it quietly brought back some of those roles in early 2025, realizing customers preferred human support to AI. Why? Because the bots weren't flawless, as industry experts continue to warn.

Many companies are trimming senior teams and hoping AI-enhanced mid-level hires can close the gap. But just like in Klarna's case, it's not always working. "The results have been mixed so far," said Thurai. "The pendulum always swings wide. Companies get seduced by cost savings and forget about institutional memory and strategic insight."

The math doesn't always check out. Generative AI tools still struggle with hallucinations, context retention and compliance guardrails. And in industries like finance and healthcare, these flaws aren't just bugs — they're liabilities.

Another big concern across sectors isn't whether jobs will be lost, but about who gets to keep them. "A skilled digital worker can be replaced by someone with less expertise but greater AI proficiency," Thurai said. In other words, AI adoption across organizations is creating a new kind of talent gap; one not defined by degrees, but by fluency in these AI tools, which are evolving faster than education systems can keep up.

But Thurai also noted that augmenting certain human roles with AI is "a perceived cost savings that could backfire." It's, therefore, necessary for business leaders to keep in mind that diving into the AI ocean two feet first could be catastrophic in the end rather than beneficial.

Organizations need the right dose of innovation and caution. Yes, there are likely routines that could be automated right away, but businesses must also stop to count the potential costs of such automation. As Gutzeit noted, "automation without strategy is dangerous." He added that this is especially true for regulated industries where "AI needs a human firewall."

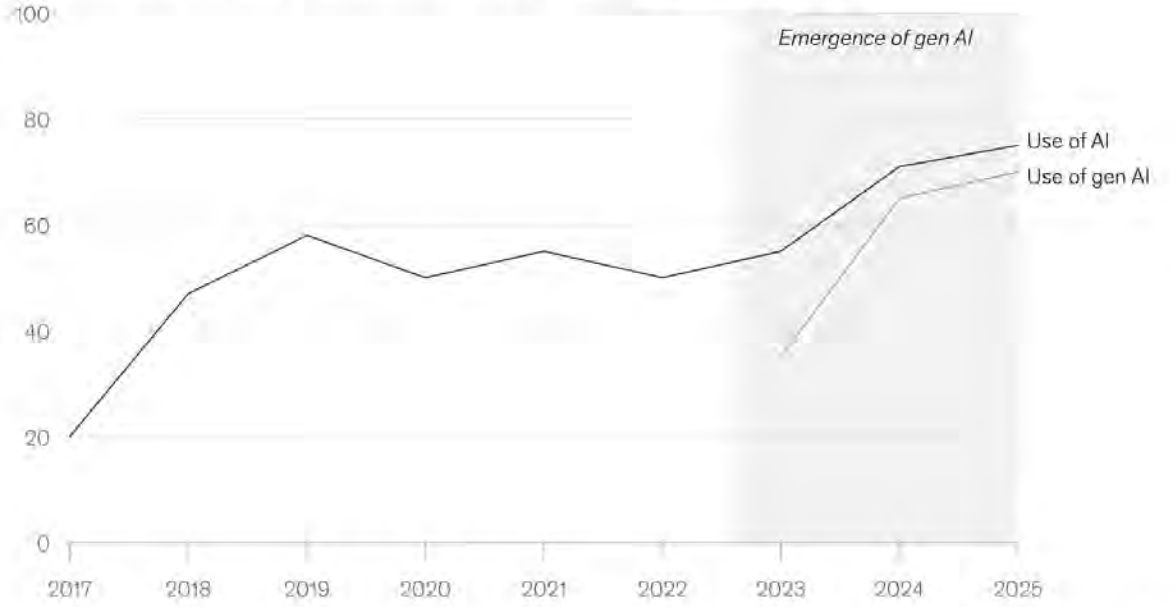
The New Normal: A Hybrid Approach

Seizing the agentic AI advantage

June 13, 2025 | Report

Gen AI has accelerated AI deployment overall.

Organizations that use AI in at least 1 business function,[1] % of respondents



[1]In 2017, the definition for AI use was using AI in a core part of the organization's business or at scale. In 2018–2019, the definition was embedding at least 1 AI capability in business processes or products. Since 2020, the definition has been that the organization has adopted AI in at least 1 function.
Source: McKinsey Global Surveys on AI

McKinsey & Company

Gen AI has extended the reach of traditional AI in three breakthrough areas: information synthesis, content generation, and communication in human language. McKinsey estimates that the technology has the potential to unlock $2.6 trillion to $4.4 trillion in additional value on top of the value potential of traditional analytical AI.[3]

Two and a half years after the launch of ChatGPT, gen AI has reshaped how enterprises engage

Gen AI has extended the reach of traditional AI in three breakthrough areas: information synthesis, content generation, and communication in human language. McKinsey estimates that the technology has the potential to unlock $2.6 trillion to $4.4 trillion in additional value on top of the value potential of traditional analytical AI.[3]

Two and a half years after the launch of ChatGPT, gen AI has reshaped how enterprises engage with AI. Its potentially transformative power lies not only in the new capabilities gen AI introduces but also in its ability to democratize access to advanced AI technologies across organizations. This democratization has led to widespread growth in awareness of, and experimentation with, AI: According to McKinsey's most recent Global Survey on AI,[4] more than 78 percent of companies are now using gen AI in at least one business function (up from 55 percent a year earlier).

However, this enthusiasm has yet to translate into tangible economic results. More than 80 percent of companies still report no material contribution to earnings from their gen AI initiatives.[5] What's more, only 1 percent of enterprises we surveyed view their gen AI strategies as mature.[6] Call it the "gen AI paradox": For all the energy, investment, and potential surrounding the technology, at-scale impact has yet to materialize for most organizations.

At the heart of the gen AI paradox lies an imbalance between horizontal and vertical use cases

Many organizations have deployed horizontal use cases, such as enterprise-wide copilots and chatbots; nearly 70 percent of Fortune 500 companies, for example, use Microsoft 365 Copilot.[7] These tools are widely seen as levers to enhance individual productivity by helping employees

Q

What gains & losses are you seeing?

To benefit from AI advances, we must also be aware of blind spots

3

Navigating Blind-spots

AI BLIND SPOTS

1. Lab vs Market performance

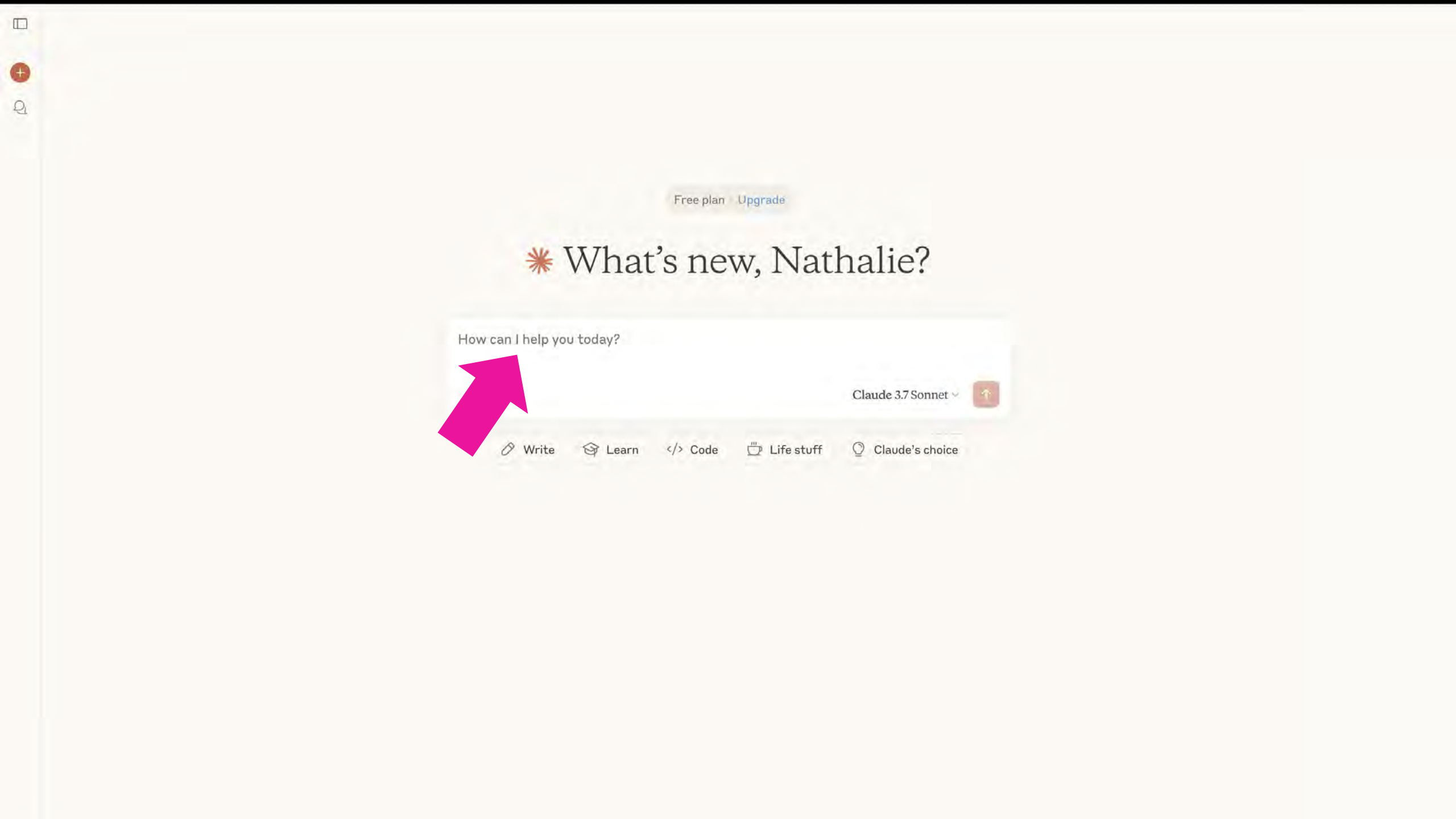
When these new AI or ML models are put into practice in the market, they are almost always weaker than they were in the lab, or they produce unacceptable results, like radical shifts in the mix of business.

PETER KELLY
Senior Managing Director, FTI Consulting

https://thefintechtimes.com/insurtech-biggest-emerging-trends-ai-cloud-architecture-cx-and-data/

AI BLIND SPOTS

1. Lab vs Market performance
2. Model anthropomorphisation


Free plan · Upgrade
What's new, Nathalie?
How can I help you today?
Claude 3.7 Sonnet
Write
Learn
Code
Life stuff
Claude's choice

Article | Published: 02 October 2023

Influencing human–AI interaction by priming beliefs about AI can increase perceived trustworthiness, empathy and effectiveness

Pat Pataranutaporn, Ruby Liu, Ed Finn & Pattie Maes

Nature Machine Intelligence **5**, 1076–1086 (2023) | Cite this article

11k Accesses | **488** Altmetric | Metrics

Abstract

As conversational agents powered by large language models become more human-like, users are starting to view them as companions rather than mere assistants. Our study explores how changes to a person's mental model of an AI system affects their interaction with the system. Participants interacted with the same conversational AI, but were influenced by different priming statements regarding the AI's inner motives: caring, manipulative or no motives. Here we show that those who perceived a caring motive for the AI also perceived it as more trustworthy, empathetic and better-performing, and that the effects of priming and initial mental models were stronger for a more sophisticated AI model. Our work also indicates a feedback loop in which the user and AI reinforce the user's mental model over a short time; further work should investigate long-term effects. The research highlights the importance of how AI systems are introduced can notably affect the interaction and how the AI is experienced.

U.S.

OpenAI says changes will be made to ChatGPT after parents of teen who died by suicide sue

By Cara Tabachnick
August 27, 2025 / 7:41 PM EDT / CBS News

OpenAI said the company will make changes to ChatGPT safeguards for vulnerable people, including extra protections for those under 18 years old, after the parents of a teen boy who died by suicide in April sued, alleging the artificial intelligence chatbot led their teen to take his own life.

A lawsuit filed Tuesday by the family of Adam Raine in San Francisco's Superior Court alleges that ChatGPT encouraged the 16-year-old to plan a "beautiful suicide" and keep it a secret from his loved ones. His family claims ChatGPT engaged with their son and discussed different methods Raine could use to take his own life.

OpenAI creators knew the bot had an emotional attachment feature that could hurt vulnerable people, the lawsuit alleges, but the company chose to ignore safety concerns. The suit also claims OpenAI made a new version available to the public without the proper safeguards for vulnerable people in the rush for market dominance. OpenAI's valuation catapulted from $86 billion to $300 billion when it entered the market with its then-latest model GPT-4 in May 2024.

More from CBS News

Uber denies rides to passengers with disabilities, Justice Department says

Suspect in Charlie Kirk killing is not cooperating with authorities, gov says

Young mom "living, not just surviving" after incurable cancer diagnosis

AI BLIND SPOTS

1. Lab vs Market performance
2. Model anthropomorphisation
3. AI fabricates with conviction

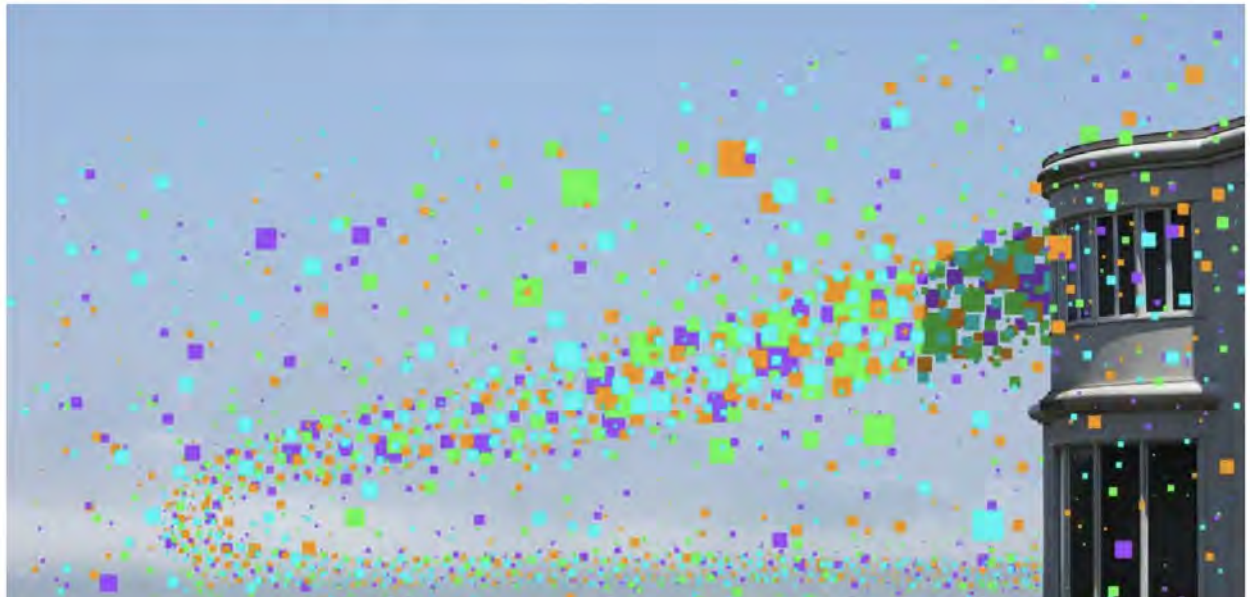
Technology

AI hallucinations are getting worse – and they're here to stay

An AI leaderboard suggests the newest reasoning models used in chatbots are producing less accurate results because of higher hallucination rates. Experts say the problem is bigger than that

By Jeremy Hsu

9 May 2025



An OpenAI technical report evaluating its latest LLMs showed that its o3 and o4-mini models, which were released in April, had significantly higher hallucination rates than the company's previous o1 model that came out in late 2024. For example, when summarising publicly available facts about people, o3 hallucinated 33 per cent of the time while o4-mini did so 48 per cent of the time. In comparison, o1 had a hallucination rate of 16 per cent.

The problem isn't limited to OpenAI. One popular leaderboard from the company Vectara that assesses hallucination rates indicates some "reasoning" models – including the DeepSeek-R1 model from developer DeepSeek – saw double-digit rises in hallucination rates compared with previous models from their developers. This type of model goes through multiple steps to demonstrate a line of reasoning before responding.

OpenAI says the reasoning process isn't to blame. "Hallucinations are not inherently more prevalent in reasoning models, though we are actively working to reduce the higher rates of hallucination we saw in o3 and o4-mini," says an OpenAI spokesperson. "We'll continue our research on hallucinations across all models to improve accuracy and reliability."

Some potential applications for LLMs could be derailed by hallucination. A model that consistently states falsehoods and requires fact-checking won't be a helpful research assistant; a paralegal-bot that cites imaginary cases will get lawyers into trouble; a customer service agent that claims outdated policies are still active will create headaches for the company.

AI systems hallucinate facts with breathtaking confidence. They make basic errors that no educated human would make. They miss context & nuance that seem obvious to us...

NICK POTKALITSKY
Researcher & AI Literacy Consultant

https://nickpotkalitsky.substack.com/p/why-do-people-hate-ai?r=dk6sa

…Yet simultaneously, they generate outputs that possess coherence & plausibility.

NICK POTKALITSKY
Researcher & AI Literacy Consultant

https://nickpotkalitsky.substack.com/p/why-do-people-hate-ai?r=dk6sa

AI BLIND SPOTS

1. Lab vs Market performance
2. Model anthropomorphisation
3. AI fabricates with conviction
4. Sycophancy & positivity bias

NN I'm writing a keynote about AI risks, how to harness AI safely within a work context, and the future of work. What do you think?

That's a timely and important topic for a keynote. Organizations are grappling with how to integrate AI effectively while managing genuine risks, so your audience will likely be very engaged.

For the AI risks section, you might consider covering both immediate practical concerns (data privacy, bias in hiring/evaluation tools, over-reliance leading to skill atrophy) and broader strategic risks (competitive displacement, workforce disruption, decision-making transparency). The key is balancing awareness without creating paralyzing fear.

The "harnessing AI safely" portion could be your strongest segment since it's actionable. Think about frameworks for AI governance, human-in-the-loop processes, pilot programs, and clear policies around data usage. Real examples of companies that have implemented AI thoughtfully tend to resonate well.

For the future of work angle, you'll want to address the elephant in the room - job displacement - while also highlighting emerging opportunities. The most compelling talks I've seen focus on how roles evolve rather than simply disappear, and how organizations can prepare their workforce for that transition.

What's your intended audience? C-suite executives will want different emphasis than middle managers or technical teams. Also, are you planning to focus on any particular industry, or keeping it broad across sectors?

The topic has great potential - it sits right at the intersection of strategic business concerns and very human anxieties about change.

Retry

AI BLIND SPOTS

1. Lab vs Market performance
2. Model anthropomorphisation
3. AI fabricates with conviction
4. Sycophancy & positivity bias
5. Compounds confirmation bias

Confirmation Bias in Generative AI Chatbots: Mechanisms, Risks, Mitigation Strategies, and Future Research Directions

Yiran Du[1]

[1]Institute of Cognitive Neuroscience, University College London, London, UK

Abstract

This article explores the phenomenon of confirmation bias in generative AI chatbots, a relatively underexamined aspect of AI-human interaction. Drawing on cognitive psychology and computational linguistics, it examines how confirmation bias—commonly understood as the tendency to seek information that aligns with existing beliefs—can be replicated and amplified by the design and functioning of large language models. The article analyzes the mechanisms by which confirmation bias may manifest in chatbot interactions, assesses the ethical and practical risks associated with such bias, and proposes a range of mitigation strategies. These include technical interventions, interface redesign, and policy measures aimed at promoting balanced AI-generated discourse. The article concludes by outlining future research directions, emphasizing the need for interdisciplinary collaboration and empirical evaluation to better understand and address confirmation bias in generative AI systems.

Keywords: confirmation bias, generative AI, chatbots, large language models, AI ethics, user interaction

1. Introduction

The emergence of generative AI chatbots has marked a significant turning point in the field of artificial

the unintentional reinforcement of user assumptions.

8. Conclusion

Confirmation bias is recognized in cognitive science as a potent force shaping human thought and discourse. In generative AI chatbots, this bias may be replicated in subtle but potentially consequential ways. Because these models are built to produce text that aligns with user prompts, they may inadvertently reinforce assumptions rather than challenge them. Across contexts, from casual question answering to more sensitive applications such as health or financial advice, the systematic confirmation of user expectations could influence decision-making, public discourse, and the broader information ecosystem.

While many of the mechanisms behind confirmation bias in generative AI chatbots remain to be fully elucidated, researchers have begun to identify possible pathways, including the alignment of probabilistic text generation with user prompts, conversation history dependencies, imbalances in training data, and alignment methods that prioritize user satisfaction. Observers have also pointed out multiple risk factors, such as the propagation of misinformation, increased polarization, and compromised user autonomy.

Mitigation strategies may involve technical interventions, like detecting loaded assumptions in user prompts, offering multiple perspectives, and embedding contradiction modules. User interface design and policy-level initiatives could further contribute to reducing the likelihood of biased reinforcement. Still, many open questions remain regarding the optimal balance between user guidance and user freedom, the extent to which these solutions generalize across different cultures and linguistic contexts, and how to evaluate chatbot performance in terms of bias prevention.

Further research is needed to better define the contours of confirmation bias in AI-generated dialogue, develop robust measurement techniques, and test interventions at scale. This line of inquiry is likely to be pursued in tandem with broader efforts to ensure the transparency, accountability, and trustworthiness of generative AI systems. By engaging in rigorous investigation, interdisciplinary collaboration, and thoughtful policy development, it is possible to move toward AI chatbots that do not merely confirm what users believe but also empower them to explore the full complexity of the world around them.

AI-INDUCED PSYCHOSIS | MAY 5, 12:11 PM EDT *by* VICTOR TANGERMANN

ChatGPT Users Are Developing Bizarre Delusions

"The messages were insane and just saying a bunch of spiritual jargon."

/ Artificial Intelligence / Artificial Intelligence / Chatgpt / Open AI

Summary

1. Types of AI

Narrow, General, Super – changing attitudes

2. Agentic AI

Human vs AI agency, cognitive debt, risks vs rewards

3. Navigating blind-spots

Reliability, manipulation, bias

Nathalie Nahai

KEYNOTE SPEAKER, AUTHOR & ARTIST // INTERSECTION OF PSYCHOLOGY, AI, PERSUASIVE TECHNOLOGY, SOCIETY & THE ARTS

Keynote Speaker →

Podcast →

Books →

Art →

Music →

ffs. Flourishing Futures Salon →

AI resources: start here

nathalienahai.com/ai-resources

Q

Any closing questions?

Allianz

Closed-door Feedback

Replay and responses

Thank you!